Photophysical characterization of isothiazologuanosine, a unique isomorphic and isofunctional fluorescent analogue of guanosine.

Olha Tkach[a,#], Lara Martinez-Fernandez[b,#], Nicolas Humbert[a,#], Ludovic Richert[a], Dmytro Dziuba[a], Pascal Didier[a], Yitzhak Tor[c], Roberto Improta[*,d], Yves Mély[*,a]

[a]Laboratoire de Bioimagerie et Pathologies, UMR 7021 CNRS Université de Strasbourg, Faculté de pharmacie 74 route du Rhin, 67401 Illkirch, France.

olha.tkach@etu.unistra.fr

nicolas.humbert@unistra.fr

ludovic.richert@unistra.fr

dmytro.dziuba@unistra.fr

pascal.didier@unistra.fr

[b] Departamento de Química, Facultad de Ciencias, Modúlo13, Universidad Autónoma de Madrid, Campus de Excelencia UAM-CSIC, Cantoblanco, 28049 Madrid, Spain.

lara.martfernandez@gmail.com

[c] Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, 92093-0358, USA.

ytor@ucsd.edu

[d] Consiglio Nazionale delle Ricerche, Istituto Biostrutture e Bioimmagini Via De Amicis 95, 80145, Napoli (Italy).

[#] Equally contributed to the work

*Corresponding authors: roberto.improta@cnr.it and yves.mely@unistra.fr

Abstract

Application of fluorescence techniques to investigate molecular interactions with nucleic acids is complicated by their poor emission, making the substitution of natural nucleobases by fluorescent nucleoside analogues (FNAs) a useful strategy. A breakthrough in fluorescent nucleoside analogues has been the development of thienoguanosine (thG) and isothiazologuanosine (^{tz}G), two isosteric mimics of guanosine. Due to its N7 atom needed in Hoogsteen base pairs and enzyme recognition, ^{tz}G is also an isofunctional G surrogate. Herein, we integrated fluorescence spectroscopy measurements with quantum mechanical (QM) calculations to characterize the mechanisms underlying ^{tz}G photophysics in different solvents. In dioxane and ethylacetate, ^{tz}G existed primarily as a H1 keto-amino tautomer with short fluorescence lifetime ($\tau \sim 2$ ns) and low quantum yield ($\phi \sim 0.10$). In buffer, the H1 tautomer ($\phi = 0.36$, $\tau = 8.84$ ns) coexisted with a weakly emissive H3 keto-amino tautomer. The two tautomers were also observed in methanol, but with a 30% decrease in ϕ and τ values for the major H1 tautomer. QM calculations suggested that the main non-radiative pathway of ^{tz}G-H1 involves NS bond loosening and is responsible for the more solvent-sensitive ϕ and τ values compared to thG. This pathway is much more efficient for ^{tz}G-H3, for which an additional pathway to dark $n\pi^*$ state and a large coupling with triplet states further explain its very low emission. This study lays the ground for rationally using ^{tz}G as a sensitive FNA .

Keywords
Ab initio calculations, fluorescent probes, fluorescence spectroscopy, nucleobases, photophysics

Introduction

Due to their outstanding sensitivity and versatility, fluorescence techniques are ideally suited for investigating biomolecules and their interactions with a wide range of ligands. Unfortunately, in the case of nucleic acids, the applications of these techniques is complicated, as the natural nucleobases are almost non-emissive,(1) frequently necessitating external labelling. Covalent labelling of oligonucleotides with bright fluorophores at their 3'- or 5'-end or at the 5-position on pyrimidines incorporated in the oligonucleotide sequence(2)(3) shows severe limitations, since the fluorophores are generally bulky and can alter the nucleic acid folding and ligand binding.(4)(5)(6) These drawbacks can be largely surpassed by using fluorescent nucleoside analogues (FNAs), such as 2-aminopurine, which can substitute natural nucleobases with limited structural or functional perturbation .(2)(7)(3)(8)(9) Although very useful, most of these FNAs do not faithfully substitute natural nucleobases, are highly quenched and/or show limited environmental sensitivity when incorporated into nucleic acids.(9) A clear breakthrough in this field was achieved by the introduction of the thieno[3,4-*d*]pyrimidine (thN) and the isothiazolo[4,3-*d*]pyrimidine (^{tz}N) families of FNAs.(10)(11) The most investigated and applied members of these families have been the analogues of G, namely thG and ^{tz}G (Figure 1).

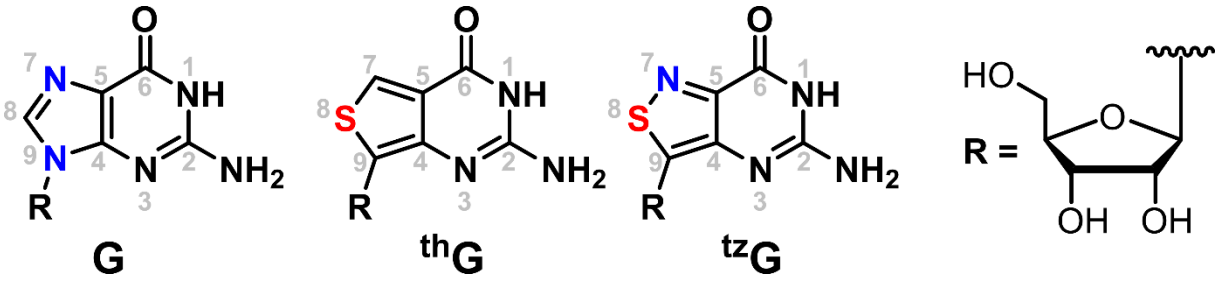

Figure 1. Structures of guanosine (G), thienoguanosine (thG) and isothiazologuanosine (^{tz}G).

A significant amount of work has been performed on thG in recent years. thG's fluorescence quantum yield (ϕ) is high (ϕ~0.50) and its dominant fluorescence lifetime in DNA is unusually long (9–29 ns), facilitating its selective measurement in complex media using time-resolved techniques. The presence of two possible tautomers with different spectral properties provides an additional information channel as their equilibrium depends on the polarity and H-bonding properties of the embedding medium. When incorporated into nucleic acids, thG retains high ϕ and environmental sensitivity, and appears to seamlessly replace G residues in DNA and RNA duplexes.(10)(11)(12)(13)(14)(15) Both in matched and mismatched duplexes, the steady-state and time-resolved fluorescence spectra of thG have been shown to sensitively report on local structural changes and single nucleotide polymorphism. Therefore, thG emerged as an outstanding emissive substitute for G.

thG differs from G by the absence of the N atom at position 9, the presence of a S atom instead of a C atom at position 8 and by the absence of the N atom in position 7. The latter feature, though not impacting the formation of canonical Watson–Crick base pairs, can impact its performance. Indeed, this N7 atom has a key role not only in the formation of Hoogsteen base pairs for the stabilization of triple helices or G-quadruplexes,(16) (17) but also in the proper recognition of G by several enzymes.(18) ^{tz}G, possessing this N7, appears, in principle, even superior to thG, being an almost fully isofunctional G surrogate. This was notably demonstrated by the much higher ability of ^{tz}G compared to thG to undergo effective deamination by guanine deaminase(18) or to be converted into a dinucleotide by a bacterial enzyme,(19) being comparable to G.

To fully exploit the potential of ^{tz}G and properly interpret its spectroscopic signatures, the first key step is to investigate its photophysics as a free nucleoside or nucleobase. This strategy has been proven to be fundamental for thG: interpreting the data of thG-labelled RNA and DNA sequences(15) (20) would have been impossible without a comprehensive characterization of the photophysical features of free thG. (10) (12) (14) (21) (22)

In contrast to thG, only preliminary spectroscopic characterization was performed on ^{tz}G,(11)(23) which has been reported to absorb with a maximum $\lambda_{abs,max}$ = 333 nm (ε = 4870 $M^{-1}\cdot cm^{-1}$) and emit at $\lambda_{em,max}$ = 459 nm (ϕ = 0.25) in water. Computational calculations(24) have suggested that the lowest-lying $^1(\pi\pi^*)$ bright state populated after light absorption evolves without any energy barrier to the lowest energy emissive minimum energy. From this structure, a high energetic barrier impedes the path to the conical intersection with the ground state, explaining the high ϕ value. However, a comprehensive time-resolved fluorescence study of free ^{tz}G as well as a study seeking for the existence of tautomers in the ground-state or excited-state is still missing. A full picture of the photophysics of ^{tz}G is not yet available, undermining its rationale use as a nucleoside surrogate in fluorescence studies of nucleic acids.

To understand the underlying photophysics of ^{tz}G, we investigated the spectroscopic properties of free ^{tz}G in two aprotic (dioxane, ethyl acetate) and two protic (aqueous buffer and methanol)

solvents, combining experimental measurements (steady-state and time-resolved fluorescence) with quantum mechanical (QM) calculations. We found that in the two aprotic solvents, ^{tz}G exhibited similar $\lambda_{em,max}$ (427 nm) and ϕ ($\approx$ 0.10). Moreover, we demonstrate that ^{tz}G mainly exists as a H1 keto-amino tautomer associated with a fluorescence lifetime of about 2 ns. Two additional minor species with lifetimes of 0.45 and 8.3 ns, respectively, were also identified. In protic solvents, ^{tz}G was found to exist as H1 and H3 keto-amino tautomers, which display different absorption and emission features. In contrast to thG, where the two tautomers show comparable brightness in aqueous buffer, the H1 keto-amino tautomer of ^{tz}G is two orders of magnitude brighter than the H3 keto-amino tautomer, exhibiting a ϕ value of 0.36 and a fluorescence lifetime of 8.84 ns. The two tautomers were also observed in methanol, but with a decreased fluorescence ϕ and lifetime for the major H1 tautomer as compared to water. This study lays the ground for the use of ^{tz}G as a sensitive fluorescence probe both in its monomeric form and when incorporated into nucleic acids.

Materials and Methods

Materials

The fluorescent isothiazologuanosine (^{tz}G) was synthesized in the form of a ribonucleoside as previously described.(11) Stock solutions of ^{tz}G (3.46 mM) were prepared in spectroscopic grade DMSO. Spectroscopic grade DMSO (99.9%), methanol (99.9%), ethyl acetate (99.5%), dioxane (99.8%) and molecular sieves (3Å) were purchased from Sigma-Aldrich. The buffer solution was 20 mM HEPES (Merck), pH 7.

UV-visible absorption and steady-state fluorescence measurements

Absorption spectra were recorded on UV-2700i (Shimadzu) and Cary 4000 (Agilent) UV-visible spectrophotometers.

Excitation and emission spectra were recorded with Fluoromax 4 and Fluoromax+ spectrofluorimeters (Jobin-Yvon) equipped with a thermostated cell compartment at 20.0 (+/- 0.2) °C. Fluorescence spectra were corrected from Raman scattering, lamp fluctuations and instrumental wavelength-dependent bias. Quartz cuvettes with 1 cm path length were used for measurements. The quantum yield (ϕ)(25) of ^{tz}G in the different solvents was calculated using quinine sulfate in sulfuric acid (0.5 M) as a reference (ϕ_R = 0.546)(26) and the following equation:

$$\phi = \phi_R \frac{I}{I_R} \frac{A_R}{A} \frac{n^2}{n_R^2} \quad (1)$$

where I is the area under the emission spectrum, A is the absorbance at the excitation wavelength, and n is the refractive index of the solvent. The quantum yields at each excitation wavelength (from 290 nm to 380 nm, each 10 nm) were given as means ± standard error of the mean for at least three independent measurements.

Time-resolved fluorescence measurements

Time-resolved fluorescence measurements were performed with the time-correlated single-photon counting (TCSPC) technique. The fluorescence intensity decays of ^{tz}G were recorded on a home-made set-up using a pulse-picked supercontinuum white-light laser (EXR-20, NKT Photonics) equipped with an UV extension (SuperK Extend UV, NKT Photonics). Excitation was at 350 and 370 nm with 19.5 MHz repetition rate. Emission was collected through a KV 389 filter and the requested wavelength was selected by a 16 mm band-pass monochromator (Jobin– Yvon H10). Photons were detected with a microchannel plate Hamamatsu R3809U photomultiplier coupled to a pulse preamplifier (HFAH 20, Becker & Hickl). Data were recorded on a TCSPC counting board (SPC 130, Becker & Hickl). The instrumental response function recorded with a polished aluminum reflector has a full width at half maximum (FWHM) of 50 ps. Alexa 488 dye with 4.1 ns lifetime was used as a reference for lifetime measurements.(27)

To rule out the possibility that the minor long-lived lifetime observed in 1,4-dioxane may be caused by the presence of H-bonding impurities, we used pre-activated molecular sieves (3Å, 4–8 mesh) to remove water from the solvent. The pre-activated molecular sieves were combined with 1,4-dioxane at a proportion of ¼ (w/w) in an oven-dried glass vial and kept sealed for 24 hours at room temperature with occasional shaking to absorb the traces of water from the solvent. Then, ^{tz}G was dissolved in the treated dioxane and time-resolved experiments were recorded immediately after.

Fluorescence lifetime data analysis

In a first step, the Maximum Entropy Method (MEM)(28)(29) was used to retrieve the distribution of the decay times without assumptions about the shape of the distribution and the number of decay components. MEM is based on maximizing the Skilling-Jane's entropy function:

$$S = \int_0^\infty \alpha(\tau) - m(\tau) - \alpha(\tau)\, log \frac{\alpha(\tau)}{m(\tau)} d\tau \qquad (2)$$

where $\alpha(\tau)$ is the reconstructed distribution, and $m(\tau)$ is the primary model. The MEM method was essentially used to provide the number of lifetime components needed to fit the decays.

To improve the determination of the decay components, a global analysis method(30)(31) was applied to analyze our multi-wavelength decay data. This method is based on the assumption that the decay components τ_i do not depend on the emission wavelength, while their amplitudes α_i are different for each wavelength. The global analysis fitting routine was built in DecayFit program(32) that uses the iterative reconvolution method and non-linear least-squares minimization method. The DecayFit program allows the use of the experimentally measured instrument response function (IRF).
Decay-associated spectra were calculated from the intensity decay curves using:(25)

$$I_i(\lambda) = \frac{\alpha_i(\lambda)\cdot\tau_i\cdot I(\lambda)}{\sum_j \alpha_j(\lambda)\cdot\tau_j} \qquad (3)$$

where I(λ) is the intensity of the steady state emission spectrum at wavelength λ.

Quantum mechanical calculations

Electronic methods. The relative energies of the different tautomers of ^{tz}G have been computed with the CBS-QB3(33) procedure, which enables an accurate determination of the thermochemical stability of medium size compounds in the gas phase (GP), and provides a benchmark for the density functionals (DFT) used for computations in solution. Two different functionals have been used in DFT and Time Dependent DFT (TD-DFT) calculations, PBE0(34)(35) and M052X,(36)(37) which have been already profitably used in our previous studies of thG and other DNA bases.(38)(39) Ground and excited state geometry optimizations, followed by the characterization of the lowest energy excited states and computation of the vertical absorption (VAEs) and emission energies (VEEs) have been performed by using different basis sets (see SI for details). Unless specified, in the main text we shall discuss the results obtained with M052X and the 6-31+G(d,p) basis set. We choose M052X as reference functional because, though overestimating the transition energies in the isolated bases,(20)(40) it has been already profitably used when studying oligonucleotides,(38)(39) since it can fairly accurately describe stacking interactions and the relative stability of charge transfer (CT) transitions,(36)(37) as shown, inter alia, by our study on thG included in DNA.(15)(20) On the average M052X overestimates the emission energies and the transition dipole moments, and these deficiencies (21)(15)(20) obviously affect the expected accuracy of the computed radiative lifetime. On the other hand, considering that vibrational effects are not included in our calculations, we here simply aim to get a qualitative estimate of the radiative lifetimes, to check if they are consistent with the experimental ones, and to obtain reference values for our future studies in oligonucleotides, where we shall use M052X. CASPT2/CASSCF calculations have been also performed, whose details can be found in the SI.

Solvent effect. Bulk solvent effects in dioxane and in water have been included in our calculations by using the Polarizable Continuum Model (PCM).(41)(42) For water, where solute/solvent Hydrogen Bonds (HB) could be important, we resorted to a hybrid implicit/explicit model, explicitly including 6 H_2O water molecules in the calculations (see below). Absorption energies have been computed at the solvent non-equilibrium level, while emission energy at the equilibrium one.

Radiative lifetimes have been estimated by using a procedure previously described,(21) based on the VEEs and the transition dipole moments. For the non-radiative decay, we have tentatively located the S_1/S_0 degeneracy regions by performing relaxed scans along relevant coordinates, knowing the limitations of TD-DFT to accurately describe such regions of the potential energy surfaces (PES).

Spin orbit coupling (SOC) terms have been computed using the pysoc program.(43)

Results and discussion

A. Photophysical properties of ^{tz}G in aprotic solvents

A1 Spectroscopic studies

A1.1 Steady-state studies

To characterize the photophysical properties of ^{tz}G, we first investigated the spectroscopic properties of ^{tz}G in 1,4-dioxane and ethyl acetate, two aprotic apolar solvents that slightly differ in their $E_T(30)$ values (36.0 and 38.1 kcal.mol^{-1}, respectively). In both solvents, the absorption

spectrum of ^{tz}G shows a maximum at $\lambda_{abs,max}$ = 341 nm. By recording the emission spectra at excitation wavelengths ranging from 300 nm to 410 nm, we found that up to 360 nm, the normalized emission spectra fully overlapped with a maximum at $\lambda_{em,max}$ = 427 nm. In contrast, at longer excitation wavelengths (370–410 nm), the emission maximum progressively shifted to 455 nm (Figure 2A, B and Figure S1A, B), indicating an additional species characterized by red-shifted absorption and emission spectra.

For greater insight, excitation spectra were recorded over the whole emission range (390–540 nm). The normalized excitation spectra largely overlapped with a maximum (340 nm) similar to the $\lambda_{abs,max}$ value, and may thus correspond to the major absorbing species (Figures 2C and S1C). However, close examination revealed a spectral broadening as the emission wavelength increased (Figures 2C and S1C, red arrow), confirming the existence of a red-shifted ground-state species. The excitation spectrum of this red-shifted species was obtained by subtracting the normalized excitation spectrum at 390 nm from the spectrum at 540 nm. Normalization to the red-part of ^{tz}G absorption spectrum provided the absorption spectrum of this red-shifted species (Figures 2A and S1A, red dashed curve and insets). Finally, subtracting this spectrum from the absorption spectrum of ^{tz}G gave the absorption spectrum of the major ground-state species (Figure 2A and S1A, blue dashed curve). The minor red-shifted species was found to be shifted by ~ 30 nm compared to the main species and contribute less than 1% to the total absorption spectrum.

From the absorption spectra, it can be further concluded that the major blue-shifted species can be selectively excited at $\lambda_{exc} \leq 320$ nm, while the red shifted species can be selectively excited at $\lambda_{exc} \geq 410$ nm. Using the emission spectra obtained at these specific λ_{exc} values (Figures 2B and S1B), we deconvoluted the emission spectrum obtained at λ_{exc} = 370 nm and found that the red-shifted species made a significant contribution in spite of its low absorbance (Figures 2D and S1D), suggesting that it exhibited a higher ϕ value than the major blue-shifted species. Both the maxima of absorbance ($\lambda_{abs,max}$ = 341 nm) and emission ($\lambda_{em,max}$ = 427 nm) of the major blue-shifted species of ^{tz}G in the two apolar solvents are comparable to the corresponding values (330 nm and 424 nm, respectively) of thG in dioxane(14), suggesting that it may correspond to the same H1 keto-amino tautomer.

The ϕ of ^{tz}G was observed to be rather constant over the 300–360 nm excitation range, being 0.10 ± 0.01 in dioxane and 0.09 ± 0.01 in ethyl acetate, respectively (Figure 3). Due to its very low absorbance, it was not possible to accurately determine the ϕ of the red-shifted species. Nevertheless, the increase of the ϕ value in dioxane at $\lambda_{exc} \geq 360$ nm confirmed that the ϕ value of the red-shifted species was higher than that of the major blue-shifted species. This increase was not observed in ethyl acetate, probably due to the very low absorbance values at $\lambda_{exc} \geq 360$ nm, that render the ϕ determination inaccurate.

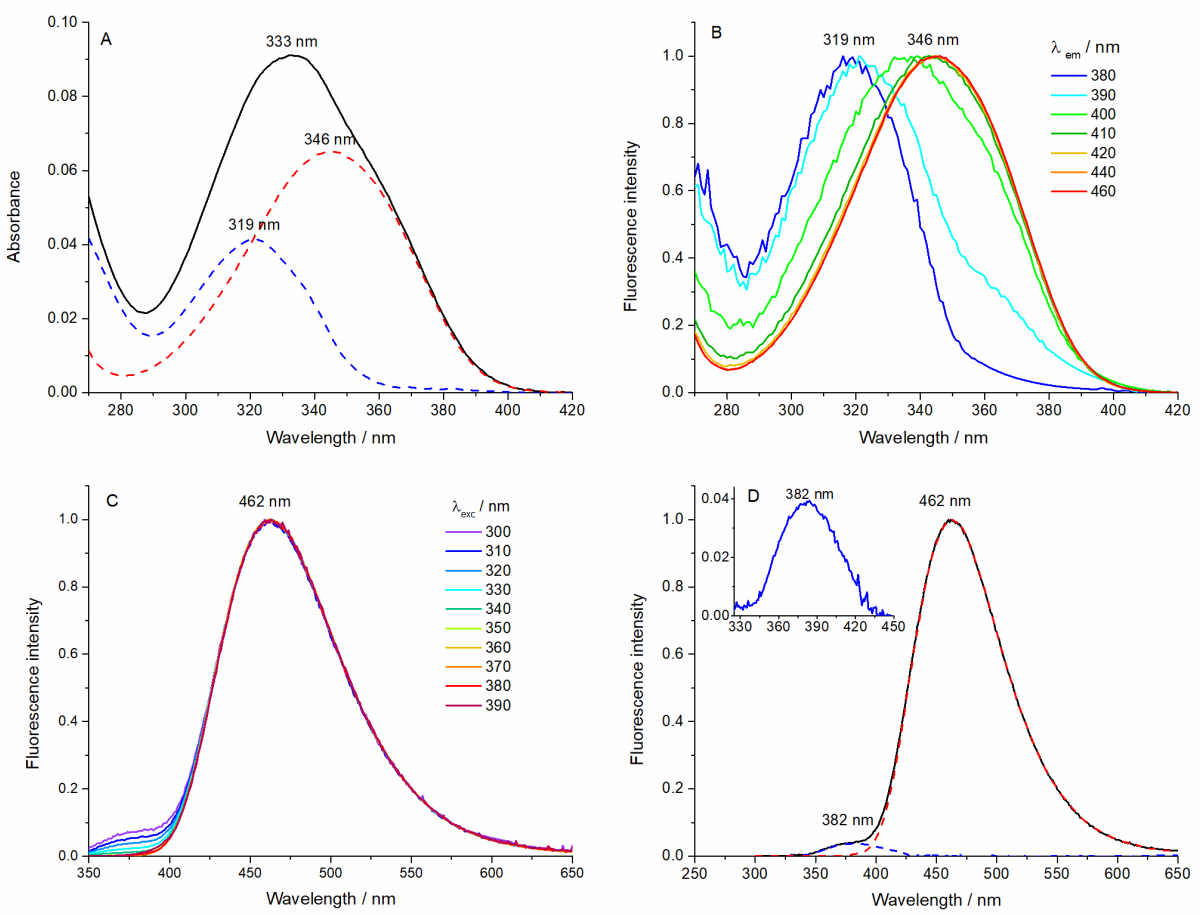

Figure 2. Absorption, emission and excitation spectra of ^{tz}G in dioxane. (A) Experimental absorption spectrum of ^{tz}G (black solid line). The spectra of the major (blue dashed line) and minor (red dashed line) ground-state species were deduced from ^{tz}G absorption and excitation spectra, as described in the main text. Inset: zoom of the absorption spectrum of the minor ground state-species. (B) Normalized emission spectra recorded at different excitation wavelengths. (C) Normalized excitation spectra recorded at different emission wavelengths. The red arrow shows the spectral broadening when λ_{em} is increased. (D) Emission spectrum of ^{tz}G recorded at λ_{exc} = 370 nm (black solid line) and its deconvolution performed by normalizing the emission spectrum of the minor red-shifted species recorded at λ_{exc} = 410 nm (red dashed line) to the red edge of the emission spectrum recorded at λ_{exc} = 370 nm. The difference between the two spectra provided the spectrum of the main emitting species (blue dashed spectrum). [^{tz}G] = 22.44 µM.

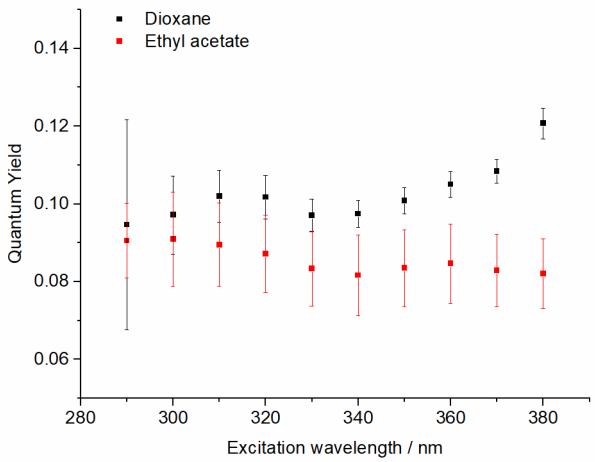

Figure 3. Dependence of ^{tz}G quantum yield on the excitation wavelength in dioxane and ethyl acetate. The graph represents mean QY values with standard errors of the means from three experiments for each solvent.

A1.2 Time-resolved fluorescence studies

Next, time-resolved experiments were performed on ^{tz}G in dioxane and ethyl acetate using λ_{exc} = 350 nm and λ_{em} ranging from 400 nm to 520 nm with 10 nm steps. The intensity decays were analyzed simultaneously by a global fit approach (Figures 4A and S2A). Analysis of the χ^2 values and distribution of residuals revealed that a three exponential function was needed for a good fit in both solvents (Figures 4B and S2B). The best fit was obtained with a short-lived lifetime (τ_1 = 0.45±0.05 ns and 0.46±0.02 ns for dioxane and ethyl acetate, respectively), a major intermediate lifetime (τ_2 = 1.97±0.09 ns and 1.73±0.05 ns for dioxane and ethyl acetate, respectively), and a long-lived lifetime (τ_3 = 8.5±0.5 ns and 8.2±0.4 ns for dioxane and ethyl acetate, respectively). No significant improvement of the χ^2 values and distribution of residuals in both solvents was observed when negative amplitudes typical of excited-state reactions were allowed. We therefore only considered the fits with positive amplitudes.

The amplitude of the major lifetime τ_2 was observed to increase with λ_{em} and reached a plateau close to 100% at $\lambda_{em} \geq 480$ nm (Figures 4C and S2C). Conversely, the amplitude of the short lifetime τ_1 dropped to zero in the red part of the emission spectrum ($\lambda_{em} > 480$ nm). The amplitude of the long-lived lifetime τ_3 increased with λ_{em} but did not exceed 2% (red squares in Figures 4C and S2C). Using λ_{exc} = 370 nm (Figure S3A-C) instead of 350 nm, the τ_3 amplitude significantly increased, reaching 6% (Figure S3C), suggesting that the τ_3 lifetime might be associated with the red-shifted minor species seen in the steady-state data. This conclusion was strengthened by the match of the decay-associated spectrum (DAS) of τ_3 with the emission spectrum of the red-shifted species (Figure S4C, inset). As the τ_3 value was similar to the major lifetime in water (see below), we checked whether it was not due to residual traces of water by adding molecular sieves for water removal. No change in the τ_3 value or amplitude was observed (data not shown), indicating that it was not an artifact due to traces of water. The DAS of the intermediate lifetime τ_2 largely dominated (> 90%) the emission of ^{tz}G in both apolar solvents (Figures S4A and B) and may thus correspond to the H1 keto-amino tautomer. Finally, the DAS of the short-lived lifetime τ_1 was found to be blue-shifted with $\lambda_{em,max}$ ~410

nm, but contributed only moderately to ^{tz}G emission. Moreover, the strong overlap of the normalized excitation spectra recorded in the blue part of the emission spectra (Figures 2C and S1C) suggested that the excitation spectrum of the species associated to τ_1 overlaps with that of the major tautomer. Both the modest contribution of this short-lived species in ^{tz}G emission and its marginal shift in respect with the major tautomer likely explain that we were unable to detect it in the steady-state spectra.

The high similarity between the steady-state and time-resolved data of ^{tz}G in dioxane and ethyl acetate strongly suggests that ^{tz}G has similar photophysical mechanisms in the two solvents. The only significant difference was the 10% decrease in both ϕ and τ_2 values in ethyl acetate compared to dioxane.

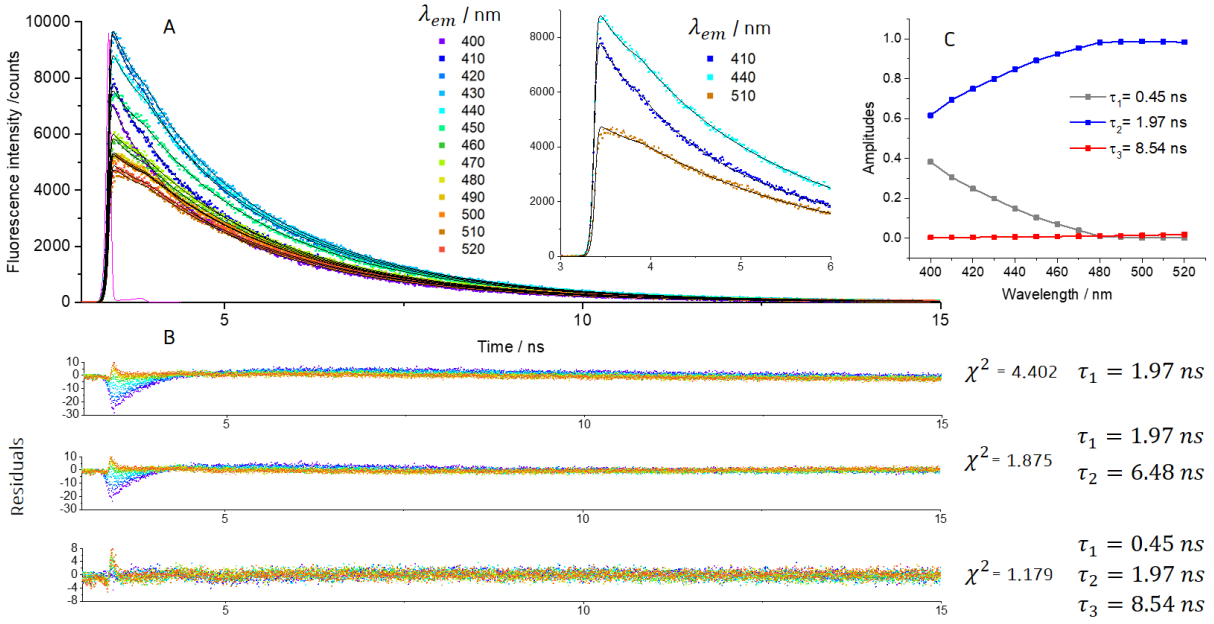

Figure 4. Time-resolved emission decays of ^{tz}G in dioxane. (A) The decays were recorded over the emission spectrum of ^{tz}G, using λ_{exc} = 350 nm. The experimental data (squares) were fitted using the global fit method (black solid lines) and a three exponential function. Experimental IRF (magenta curve). Inset: highlight of the intensity decays and their fits in the first channels for three λ_{em} values. (B) Distribution of residuals, χ2 values and lifetime values for one, two, and three exponentials fit of the decays. (C) Wavelength dependence of the amplitudes of the lifetime components obtained using a three exponential fit. [^{tz}G] = 22 µM.

A2. Quantum mechanical calculations

A2.1 Tautomerization equilibria

To identify the species responsible of the observed spectroscopic properties, we resorted to QM calculations, focusing both on tautomeric and conformational equilibria. To reduce the associated computational cost, most analyses were performed on a computational model, where the ribose unit is mimicked by a methyl group (Figure 5). We nevertheless also performed test calculations on the ribonucleoside, focusing on the syn/anti conformational equilibrium of aromatic heterocycle with respect to the tetrahydrofuran ring (Figure S5), which could impact the photophysics of ^{tz}G in non-polar solvents.

As a first step of our analysis, we compared the relative energy (Table 1) of the different tautomers (Figure 5), and the dependence on the nature of the embedding medium. In the gas phase, we have checked the reliability of DFT calculations (including different functionals) and basis sets, by using the CBS-QB3 approach. As shown in Table 1, ^{tz}G-H1 is predicted to be the most stable tautomer. M052X and PBE0 estimates are in good agreement with those provided by CBS-QB3, the discrepancy being, with a few exceptions, smaller than 1 kcal/mol.

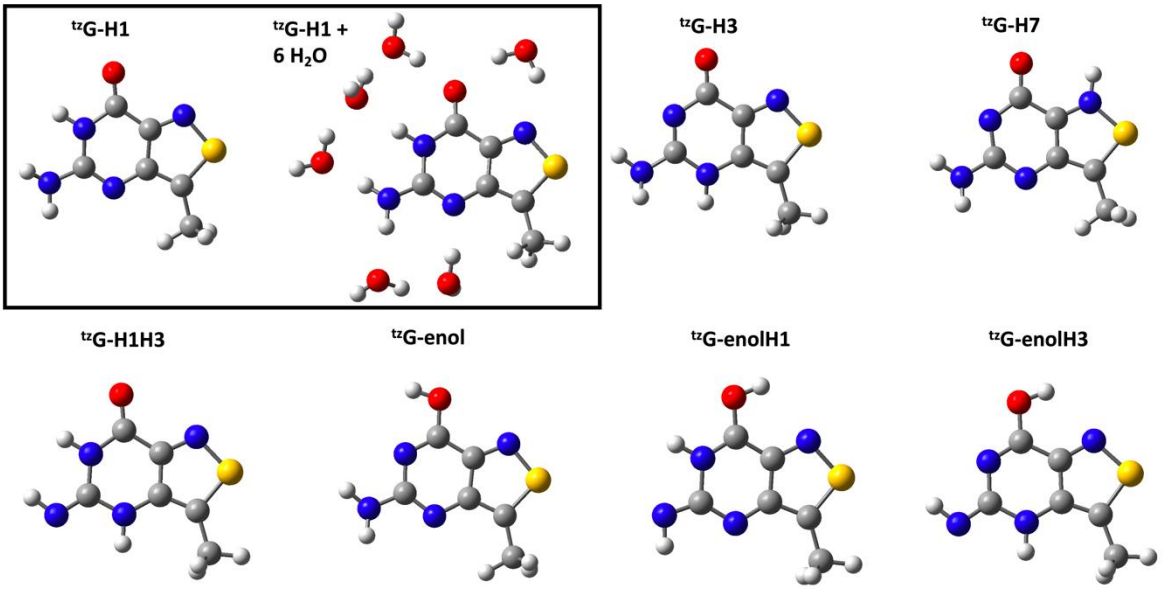

Figure 5. Schematic drawing of the main tautomers of ^{tz}G with an example of the explicit/implicit solvent model for the ^{tz}G-H1 tautomer.

Table 1: Relative stability (ΔG in kcal/mol) of the main tautomers of ^{tz}G according to different computational methods and embedding media

	CBS-QB3 Free Energy	**M052X/ 6-31+G(d,p)**	**PBE0/ 6-31+G(d,p)**	**M052X/ 6-31G(d)**	**PBE0/ 6-31+G(d)**
Gas Phase					
^{tz}G-H1	0.0	0	0	0	0
^{tz}G-H3	9.8	10.7	10.3	10.2	10.6
^{tz}G-H7	23.9	23.7	21.3	25.4	24.2
^{tz}G-Enol	3.4	2.9	3.9	4.9	5.9
^{tz}G-H1H3	7.3	8.1	7.6	7.9	7.3
^{tz}G-EnolH1	31.6	33.1	32.0	34.9	34.0
^{tz}G-EnolH3	15.9	16.3	17.0	17.9	18.3
Dioxane					
^{tz}G-H1		0.0	0.0	0.0	0.0

^{tz}G-H3		7.2	8.1	8.8	8.6
^{tz}G-H7		19.9	20.1	24.5	22.4
^{tz}G-Enol		4.3	5.9	6.7	7.6
^{tz}G-H1H3		7.4	8.1	8.3	7.8
^{tz}G-EnolH1		32.9	33.0	35.9	35.2
^{tz}G-EnolH3		16.7	18.2	19.1	19.8
6 H_2O					
^{tz}G-H1		0.0 (93%)		0.0	
^{tz}G-H3		1.6 (7%)		2.8	
^{tz}G-H7		12.8 (0%)		12.9	
^{tz}G-Enol		6.4 (0%)		11.35	

Once the accuracy of our DFT estimates was confirmed, we have used this method to study the tautomeric equilibria in dioxane, where the solvent effect has been considered by the PCM model. The relative stability of ^{tz}G-H3 in dioxane is higher with respect to the gas phase, whereas that of the ^{tz}G-Enol tautomer is slightly lower (Table 1). Moreover, our calculations suggest that ^{tz}G-H1 should be the only ground state tautomer present in dioxane and thus, the main species observed in the excitation and emission spectra. The only other tautomer which could potentially exist is ^{tz}G-Enol, but its contribution should be less than 1%. This conclusion is confirmed by our test calculations considering the ribose ring. As reported in Table S1, the most stable conformer is ^{tz}G-H1$_{anti}$, which is close in energy with the syn one, and 4.1 kcal/mol more stable than ^{tz}G-Enol$_{anti}$ one, which is 0.8 kcal/mol more stable than ^{tz}G-Enol$_{syn}$.

A2.2 Photophysics

The lowest energy excited state (S_1) at Franck Condon (FC) for ^{tz}G-H1 tautomer is associated with a large oscillator strength (f = 0.23) and can be described as a ππ* excitation with a predominant HOMO→LUMO character (Table 2). This transition, labelled in the following as ππ*1, is associated to the lowest energy absorption band of ^{tz}G in dioxane. Quantitatively, the computed Vertical Absorption Energy (4.33 eV, 286 nm) is blue shifted by 0.69 eV with respect to the maximum of the lowest energy band of the experimental spectrum (3.64 eV, 341 nm). This discrepancy is partially due to the lack of vibrational effects,(44) whereas the rest depends on the inaccuracy of the chosen computational approach (see Computational details). The S_2, ~0.5 eV higher in energy, corresponds instead to a dark nπ* state (f = 0.00) which can be described as an excitation from the Lone Pair (LP) of the carbonyl oxygen atom (with some contribution of the "N7" LP) to the LUMO π*. Therefore, we label this excited state as nπ*1. Geometry optimizations of S_1 leads to a stable and strongly emissive S_1 minimum (Vertical Emission Energy=3.27 eV, 429 nm), where the ring retains an almost planar geometry. The computed Stokes shift for the emission from this minimum is 1 eV, ~90 nm, a value in reasonable agreement, considering the lack of vibronic effects, with the experimental one (0.74 eV), supporting the attribution of the main emitting species to the H1 tautomer.

Table 2. Main photophysical properties of ^{tz}G tautomers in dioxane. Adiabatic energies (eV) with respect to the corresponding S_0 minimum and associated wavelengths (nm) in parentheses. Emission energies (eV) and corresponding wavelengths (nm) are in italics. PCM/TD-M052X/6-31+G(d,p)//M052X/6-31+G(d,p) level of theory. Oscillator strengths (f) are given in brackets. The adiabatic energies of the excited state minima with respect to the S_0 minimum of ^{tz}G-H1 are in bold, while the associated wavelengths are in parentheses.

	^{tz}G-H1	^{tz}G-H3	^{tz}G-H7	^{tz}G-Enol
		Dioxane		
S_1	4.33[0.23] (286)	4.20[0.00] (295)	3.25[0.15] (381)	3.94[0.21] (315)
S_1 minimum	**3.86**[0.15]*3.27* (379) *429*	**4.22**[0.00]*3.45* (293) *359*	**3.81**[0.10]*2.47* (325)*502*	**3.73**[0.18]*3.14* (394)*455*
S_2	4.82[0.00]	4.70[0.30]	3.64[0.00]	4.58[0.00]
S_2 minimum	**4.41[**0.01]*3.89*[a] (281) *319*	**4.60** [0.06] *2.98* (270) *416*	Decay to ππ*	Decay to ππ*
S_3	5.22[0.00]	5.35[0.00]	4.28[0.00]	5.21[0.17]
S_4	5.47[0.01]	5.53[0.32]	4.29[0.12]	5.62[0.00]
S_5	5.71[0.00]	5.73[0.00]	4.41[0.00]	5.64[0.01]
S_6	5.83[0.76]	5.80[0.00]	5.66[0.01]	5.99 [0.59]

[a]Single point on the minimum optimized at the 6-31G(d) level, since at the 6-31+G(d,p) level a crossing with S1 is always found

The lowest energy FC excited states for the other tautomers are similar (Table 2) to those described for ^{tz}G-H1, with a bright ππ∗ state and a dark LP nπ* state involving the carbonyl oxygen and/or the nitrogen atoms. However, the energies of ππ*1 and nπ*1 and, therefore, their relative stability strongly depend on the tautomer. For ^{tz}G-H3, the dark nπ*1 state (S_1, 4.20 eV) is more stable than ππ*1 (S_2, 4.70 eV) (Table 2). Moreover, in the FC region the ππ*1 transition in ^{tz}G-H7 and ^{tz}G-Enol is significantly red-shifted with respect to ^{tz}G-H1. It is thus not surprising that when optimizing the geometry of ππ*1 for the ^{tz}G-H7 and ^{tz}G-Enol, the resulting minimum is more stable than that of ^{tz}G-H1, by ~0.1 eV for both the ^{tz}G-Enol and ^{tz}G-H7 (see their absolute energy with respect to the ^{tz}G-H1 S_0 minimum, reported in bold in Table 2).

Interestingly, the computed Vertical Absorption Energy (3.94 eV, 315 nm) of ^{tz}G-enol is red shifted by ~50 nm (0.6 eV) with respect to the experimental absorption maximum (3.35 eV, 369 nm) of the minor red-shifted absorption species. In addition, a 60 nm (0.4 eV) shift can be seen between the computed Vertical Emission Energy (3.14 eV, 394 nm) of ^{tz}G-enol and the experimental emission maximum (2.72 eV, 455 nm) of the minor red-shifted absorption species. Since both shifts are comparable to those observed for the major H1 tautomer, the minor red-shifted species may be attributed to the ^{tz}G-enol tautomer. This assignment is also supported by the agreement between the calculated (Table 1) and experimental (Figure 2A) percentages of this tautomer in the ground state. This picture also holds when the ribose is

included in the computational model (Table S1). However, as detailed in the SI, the emissive properties of the minima show some dependence on the syn/anti equilibrium of the ring, which affects the planarity of the minima (Figure S6).

Finally, we have estimated the fluorescence radiative lifetime (τ_{rad}^{comp}) of the H1 tautomer, based on State Specific PMC-TD-M052X calculations, which should more accurately describe the emission energies. We obtained τ_{rad}^{comp} =15.8 ns, fully consistent with the experimental value of $\tau_{rad} = \tau_2/\phi$ = ~20 ns, considering that this tautomer is by far, the main emitting species.

Non radiative decay paths. We next characterized the main non-radiative decay channels of the lowest energy bright state. The main paths involve (i) decay to the ground electronic state (S_0) through a conical intersection (CI), (ii) population transfer to a dark, non-emissive singlet electronic state and (iii) population transfer to a triplet electronic state.

For the first path, considering that TD-DFT is not the most suitable method to locate the CI with the ground electronic state, we integrated our analysis with CASPT2 calculations to provide an additional check for the TD-DFT picture. CASPT2 predicts that the lowest energy S_1/S_0 CI involves the loosening of the N7-S bond, leading to what in the following we label as Ring-Opening CI (Ring Op-CI in Figure 6 and SI). In addition to this degeneracy region, we explored two additional possible crossing regions between $\pi\pi^*1$ and S_0. One, labelled oopS in the following, is characterized by an out of plane S atom and resembles the CI reported for thG.(21) The third one, marked as nh2, displaces the C2 and its NH_2 substituent out of the plane and is the well-known lowest energy CI for guanine.(45) As discussed in the SI, the two most effective decay channels involve the oopS-CI and the RingOP-CI, the latter being characterized by a vanishingly small energy barrier. After the crossing with the S_0 surface, the system recovers its minimum, reforming the initial NS bond, in line with the photostability of ^{tz}G (data not shown). M052X provides a picture very close to that of CASPT2, supporting the reliability of the analysis we have performed for H1 and enol tautomers in dioxane.

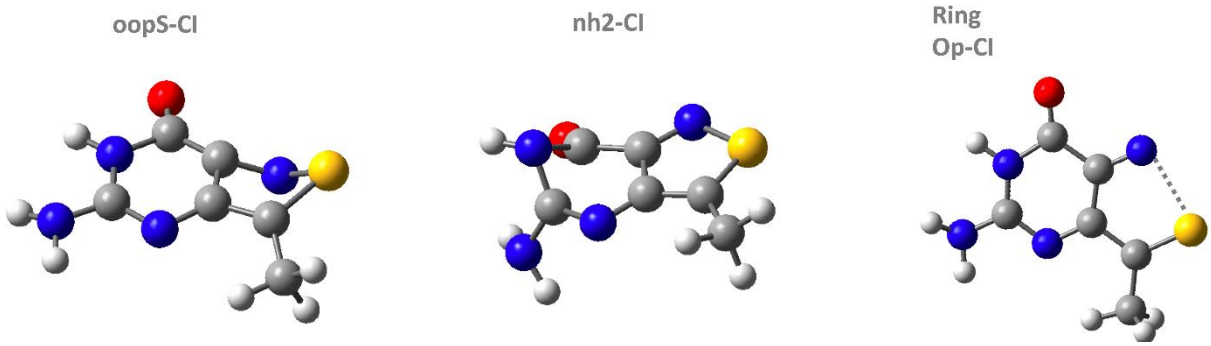

Figure 6. Three different S_1/S_0 potential crossing regions investigated in this study.

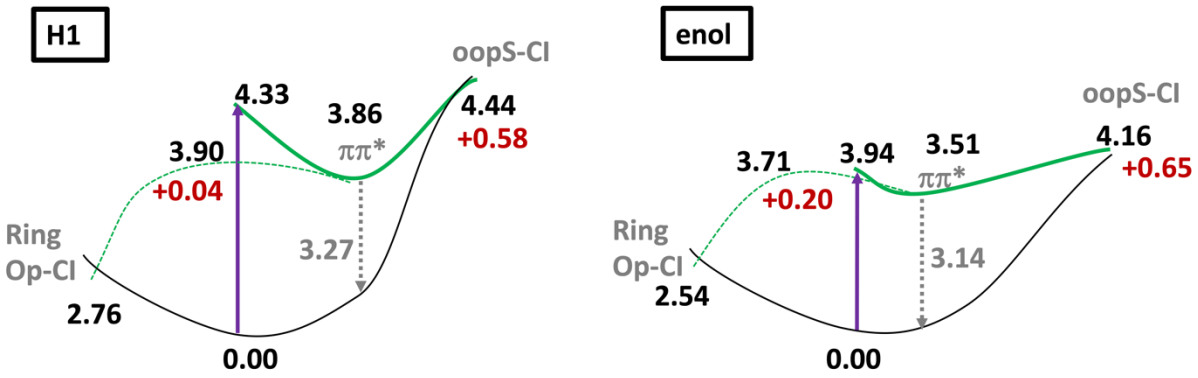

Figure 7. Potential Energy Surfaces in dioxane. PCM/TD-M052X/6-31+G(d,p) calculations. Adiabatic Energies in black, vertical energies in gray and barriers in red (eV). The dashed green curve schematically depicts the path connecting the minima of the ππ* state with the Ring-Op-CI.

As depicted in Figure 7, in dioxane we obtain a picture similar to that found in the gas phase (Figures S7 and S8). A quite substantial energy barrier (0.6 eV) separates the minima of the ππ* state from the oopS crossing region. In contrast, the access to the RingOp region is extremely easy, especially for H1 tautomer, for which the energy barrier is 0.04 eV. This value increases to 0.07 eV, when the larger aug-cc-pvtz basis set is used. Interestingly, the energy barrier towards Ring Op-CI region is larger for the enol tautomer (0.25 eV at the aug-cc-pvtz level), which is consistent with our proposal that this tautomer is more emissive. For the second path, for both the H1 and enol tautomers, the population transfer to the closest nπ* state is expected to be marginal, considering the large energy gap (> 0.5 eV) existing in the FC region. Concerning the third path related to triplet population, the H1 tautomer has two lowest lying triplet states with modest spin orbit coupling (SOC < 5 cm^{-1}) at the position of the ππ*1 minimum. In contrast, the close lying state T3 (+0.6 eV) shows significant SOC (24 cm^{-1}). The energy gap with the triplet states is similar for the enol minimum but with smaller SOC values (< 1 cm^{-1} with T_1 and T_2 and 16 cm^{-1} with T_3).

Summarizing the experimental and computational analyses, it can be inferred that the keto-amino H1 tautomer is the major tautomer in dioxane. It exhibits an emission maximum at 428 nm. The relatively low ϕ (0.10) of ^{tz}G in dioxane can be explained for the H1 tautomer by the existence of an easily accessible path to the ground state, involving the loosening of the NS bond. The minor red-shifted tautomer with an emission maximum at about 455 nm was identified as the enol tautomer. Our calculations do not provide any definitive indication for the identity of the short-lived lifetime observed in the time-resolved data. One possible explanation is related to the different sugar conformers (syn/anti, without considering the different possible tetrahydrofuran ring conformations) that have slightly different photophysical properties and could display different lifetimes. Alternatively, the different regions of the PES (e.g. the non-planar structures with respect to the planar ones) may have slightly different photophysical properties. As a result, a wave-packet exploring different regions of the PES could lead to different lifetimes.

B. Photophysical properties of ^{tz}G in protic solvents

B.1 Spectroscopic studies

B1.1 Steady-state fluorescence studies

The absorption spectrum of ^{tz}G in aqueous buffer at pH 7 shows a maximum at $\lambda_{abs,max}$ = 333 nm, but is clearly not symmetric, exhibiting a shoulder in its red edge (Figure 8A). In line with this observation, the excitation spectra recorded at different emission wavelengths showed a progressive shift in their maximum from 319 nm to 346 nm (Figure 8B), indicating that multiple species coexist in the ground state. By analogy with the parent thG nucleoside, we hypothesized that there are two tautomers in equilibrium. Since the red-shifted excitation spectra fully overlapped for $\lambda_{em} \geq 420$ nm, we assumed that they corresponded to the spectrum of one of the two tautomers and used it for deconvoluting the absorption spectrum. By subtracting the red-shifted excitation spectrum normalized to the red-edge part of the absorption spectrum, we obtained a blue-shifted band with $\lambda_{abs,max}$ = 319 nm (Figure 8A) that overlapped well with the excitation spectrum recorded at λ_{em} = 380 nm (Figure 8B). The absorption maxima of the two ^{tz}G tautomers [$\lambda_{abs,max}$ = 319 nm and 346 nm (Figure 8A)] corresponded quite well to the maxima of the two previously characterized thG tautomers ($\lambda_{abs,max}$ = 313 and 334 nm),(14) suggesting that they could correspond to the H3 and H1 keto-amino tautomers, respectively. By assuming that the molar absorption coefficient of the H1 tautomer in buffer at pH 7 is similar to that in dioxane (4650 $M^{-1}cm^{-1}$)(11) where the H1 tautomer is largely dominant (see above), we calculated the concentration of the H1 tautomer from the deconvoluted absorption spectrum (Figure 8A) and thus, deduced the concentration and molar absorption coefficient for the H3 tautomer (9196 $M^{-1}cm^{-1}$). Based on these data, the H1/H3 concentration ratio was inferred to be 3/1. Comparison of the absorption spectra of the two tautomers (Figure 8A) further revealed that the H1 tautomer can be selectively excited at wavelengths ≥ 360 nm.

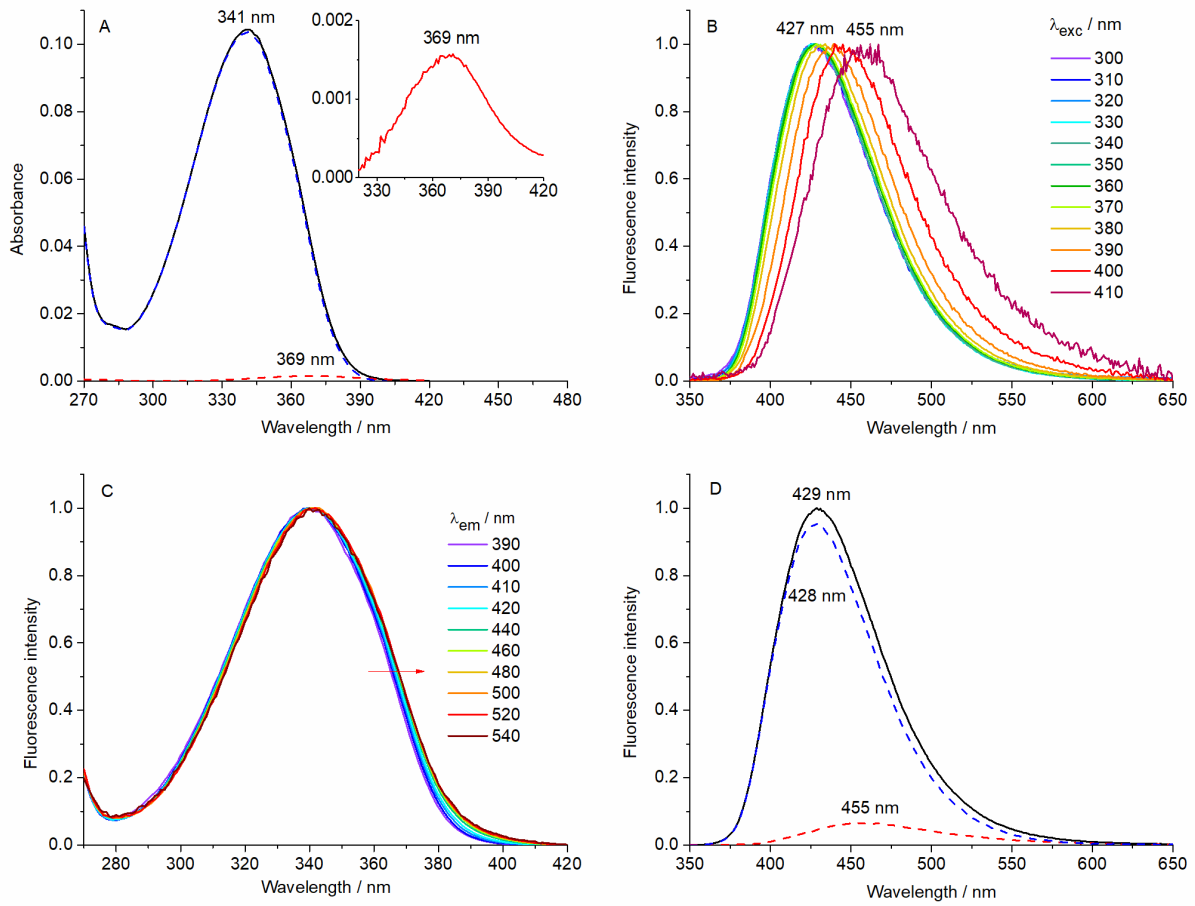

Figure 8. Absorption, excitation and emission spectra of ^{tz}G in Hepes 20 mM buffer pH 7.0. (A) Absorption spectrum of ^{tz}G (black solid line) and its deconvolution (dashed lines). The red-shifted spectrum (red dashed line) was obtained by normalizing the excitation spectrum obtained at λ_{em} = 460 nm to the red edge of the absorption spectrum. The blue-shifted spectrum (blue dashed line) was obtained by subtracting the red-shifted spectrum from the experimental absorption spectrum. (B) Normalized excitation spectra recorded at different emission wavelengths. (C) Normalized emission spectra recorded at different excitation wavelengths. (D) Emission spectra of ^{tz}G recorded at λ_{exc}= 320 nm (black line) and 380 nm (red dashed line). The two spectra were normalized at their emission maximum (462 nm). The small blue-shifted spectrum (blue dashed line) with $\lambda_{em,max}$ = 382 nm was obtained by subtracting the red spectrum from the black one. Inset: highlight of the blue-shifted emission spectrum. [^{tz}G] = 18.72 µM.

Emission spectra recorded at $\lambda_{exc} \geq$ 360 nm provided a single emission band centered at $\lambda_{em,max}$ = 462 nm (Figure 8C), close to the $\lambda_{em,max}$ value (470 nm) of the thG H1 tautomer, strengthening the assignment of the red-shifted absorption band to the H1 tautomer. In contrast to thG, excitation at λ_{exc} < 360 nm did not change the emission maximum, but led to a small shoulder with $\lambda_{em,max}$ = 382 nm, close to that of thG H3 tautomer (400 nm). The intensity of this small peak was the highest at λ_{exc} =320 nm, where absorbance of the H3 tautomer is the highest. This small peak was, thus, attributed to the emission of the H3 tautomer. The low emission of this tautomer is in variance with thG, where the H3 and H1 tautomers have similar brightness values.

The ϕ value was observed to increase with λ_{exc} in the range 280–360 nm, reaching a constant value of 0.36 ± 0.01 at $\lambda_{exc} \geq 360$ nm (Figure 9). Since the absorption of the blue tautomer is negligible above 360 nm, this value corresponds to the ϕ of the H1 tautomer. A ϕ value of less than 0.01 was estimated for the H3 tautomer from the inset of Fig 8D, using 2-aminopurine (2AP) as a standard (ϕ=0.68). While both thG tautomers have similar ϕ (0.50), the ϕ values of the two ^{tz}G tautomers differ by two orders of magnitude.

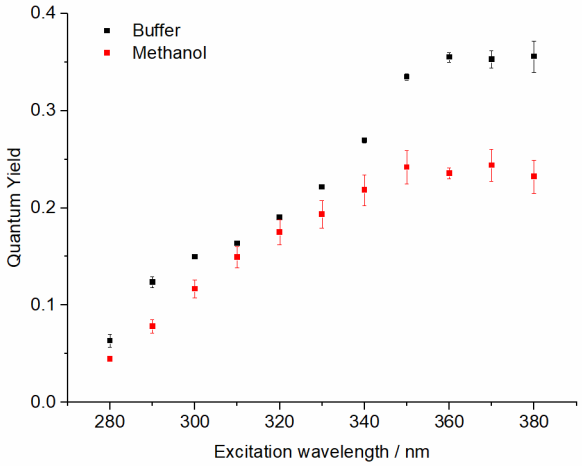

Figure 9. ϕ values of ^{tz}G in buffer and methanol at different excitation wavelengths. The ϕ values are represented as means ± standard error of the mean for three experiments in Hepes buffer, pH 7.0 (black squares) and two experiments in methanol (red squares).

The spectroscopic properties of ^{tz}G were also characterized in methanol (Figure S9). As in buffer, the deconvolution of the absorption spectrum (Figure S9A) and the excitation spectra (Figure S9B) highlighted the presence of two ground-state species with similar absorption/excitation maxima to those observed in buffer. Consequently, the two species were attributed to H1 and H3 tautomers (Figure S9A). Assuming that the molar absorption coefficient of the H1 tautomer in methanol is similar to that in dioxane (ε = 4650 $M^{-1}cm^{-1}$)(11), the calculated percentage of ^{tz}G-H3 in methanol was decreased by half compared to buffer, representing only 13% of the total species present in solution. This decrease of the H3 tautomer percentage with solvent proticity suggested that, similar to thG, the H3 tautomer of ^{tz}G is stabilized by H-bonding. Deconvolution of the emission spectra of ^{tz}G in methanol (Figure S9C) showed, as in aqueous buffer, a major contribution of the H1 tautomer centered at $\lambda_{em,max}$ = 455 nm and a very minor contribution of the H3 tautomer centered at $\lambda_{em,max}$ = 382 nm (Figure S9D). The emission maxima of ^{tz}G H1 and H3 tautomers are close to those of thG-H1 (458 nm) and thG-H3 (390 nm) tautomers in methanol. As expected from the environmental sensitivity of ^{tz}G, the emission of the H1 tautomer was blue-shifted by 7 nm for methanol as compared to buffer, indicating positive solvatochromism. As for buffer, the ϕ of ^{tz}G in methanol (Figure 9) grew with λ_{exc} and reached a constant value (0.24 ± 0.01) at $\lambda_{exc} \geq 360$ nm, which thus corresponds to the ϕ of the H1 tautomer. Due to its very low emission, the ϕ value of the H3 tautomer could not be accurately determined but was estimated to be below 0.01.

B.1.2 Time-resolved fluorescence studies

Time-resolved experiments were performed on ^{tz}G in buffer and methanol using λ_{exc} = 350 nm and λ_{em} ranging from 410 nm to 550 nm (Figure 10A, Figure S10A). The χ^2 values and distribution of residuals showed that a two exponential function provided an optimal global fit of the intensity decays of ^{tz}G in buffer and methanol (Figure 10B, Figure S10B). Optimal fits were obtained with lifetimes of 0.27(±0.06) and 8.84(±0.26) ns in water, and 0.28(±0.06) and 6.12(±0.39) ns in methanol. In both media, the amplitude of the long-lived lifetime τ_2 increased with λ_{em}, reaching a plateau at $\lambda_{em} \geq 480$ nm (Figures 10C and S10C). Conversely, the amplitude of the short-lived lifetime τ_1 dropped to zero at $\lambda_{em} \geq 480$ nm. Since at λ_{exc} = 350 nm, mainly the H1 tautomer absorbs, τ_2 can be attributed to the H1 species with higher ϕ. The value of τ_2 is shorter by 1/3 in methanol, which matches with the decrease in ϕ of the H1 tautomer in methanol as compared to buffer. The τ_1 component is mainly present in the blue part of the emission spectrum, and can thus be associated to the H3 species excited to a small extent at λ_{exc} = 350 nm (Figure 8A). Both in buffer and methanol, the DAS of τ_2 almost completely overlapped with the steady-state spectrum (Figure S11A and B), in agreement with the deconvolution of the emission spectra in Figure 8D and S9D. The DAS of the τ_1 component was too negligibly small to allow a proper comparison with the spectrum of the H3 tautomer in the inset of Figure 8D and S9D. Nevertheless, its blue shift and its fluorescence lifetime (nearly two orders of magnitude shorter than that of the H1 tautomer) are fully consistent with the assignment of this spectrum to the H3 tautomer.

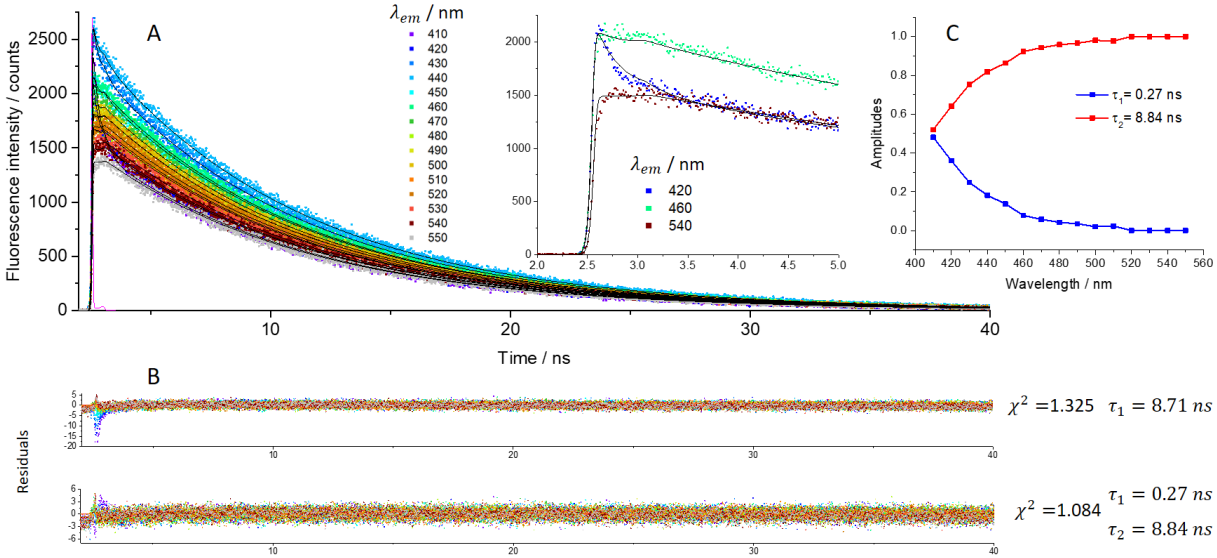

Figure 10. Time-resolved emission decays of ^{tz}G in buffer. (A) The decays were recorded over the emission spectrum of ^{tz}G, using λ_{exc} = 350 nm. The experimental data (squares) were fitted by the global fit method using a two exponential function (black lines). The experimental IRF (magenta line) was used for the fit. Inset: highlight of the intensity decays and their fits in the first channels for three λ_{em} values. (B) Distribution of residuals, χ2 and lifetime values for a one and two exponentials fit. While the χ2 value is only slightly improved by using a two exponential fit as compared to a one exponential fit, a significant improvement in the distribution of the residuals was observed especially, for the first channels. (C) Wavelength dependence of the amplitudes of the fluorescence lifetimes obtained by a two-exponential fit of the decays in (A). [^{tz}G] = 20.5 µM, buffer : Hepes 20 mM at pH = 7.

B.2. Quantum mechanical calculations

B2.1 Tautomerization equilibria in the ground state

The H1 tautomer of ^{tz}G is predicted to be the most stable also in water (Table 1). However, in this medium ^{tz}G-H3 exhibits a similar stability, being only 1.5 kcal/mol less stable at the M052X/6-31+G(d,p) level. Considering that, at this level of theory, the relative stability of ^{tz}G-H3 tautomer is underestimated by 0.9 kcal/mol with respect to CBS-QB3 in the gas phase, we can estimate that the energy difference between the two tautomers is < 1 kcal/mol, making the population of ^{tz}G-H3 in water significant, in full agreement with the experimental data.

B2.2 Photophysics

In water (Table 3), the lowest energy excited states in the FC region are qualitatively very similar to those described in dioxane. The main difference concerns the relative stability of nπ*1 for ^{tz}G-H3, whose excitation energy blue shifts to 4.90 eV (Table 3). Confirming the trends known for the natural nucleobases,(38)(39) this nπ* transition is thus destabilized by solute-solvent hydrogen bonds. The minima optimized for the different tautomers in water are similar to those found in dioxane, but the computed emission energies exhibit significant solvent shifts. Interestingly, the emission maximum computed for the most stable tautomer (^{tz}G-H1) is red-shifted in water with respect to dioxane, in agreement with the experimental data. Moreover, our calculations indicate that the emission of the minor tautomer (^{tz}G-H3) is blue-shifted by ~0.5 eV with respect to that of ^{tz}G-H1, supporting the hypothesis that the small shoulder in the fluorescence spectrum in water is associated with ^{tz}G-H3.

Table 3. Main photophysical properties of ^{tz}G tautomers in water solution. Adiabatic energies (in eV) with respect to the corresponding S_0 minimum. Emission energies (in eV) in red/italics. PCM/TD-M052X/6-31+G(d,p)//M052X/6-31+G(d,p) level of theory. The wavelengths in parentheses were calculated from the energies. The wavelengths in green are the experimental wavelengths. Oscillator strengths are given in brackets.

	^{tz}G-H1	^{tz}G-H3	^{tz}G-H7	^{tz}G-Enol
		H_2O		
S1	4.06[0.19] (305nm)346nm	4.52[0.27] (274 nm)319nm	3.58[0.19] (346 nm)	3.88[0.19] (320 nm)
S1 minimum	3.53[0.25]*3.11* *(399 nm) 462 nm*	4.08[0.29]*3.51* *(352 nm) 382 nm*	3.73[0.21]*2.74* *(452 nm)*	3.69[0.25]*2.96* *(419 nm)*
S2	5.08[0.00]	4.90[0.00]	4.38[0.00]	4.80[0.00]
S2 minimum		4.51(0.00)*3.57*		
S3	5.39[0.03]	5.39(0.32]	4.80[0.15]	5.31[0.16]

Resorting to State Specific PMC-TD-M052X calculations, we have estimated the fluorescence radiative lifetime (τ_{rad}^{comp}) of H1 and H3, the most populated tautomers in water. We have obtained τ_{rad}^{comp} =12.7 ns and 7.3 ns for H1 and H3, respectively. Considering that the M052X

computed emission energies are overestimated with respect to the experimental ones, the value estimated for H1 is in a reasonable agreement with the experimental one, $\tau_{rad} = \tau_2/\phi$ = ~24 ns.

Non radiative decay paths. We have mapped the main non-radiative decay paths, focusing on the H1 and H3 tautomers populated in water. For the decay of H1 to S_0 we obtain a pathway close to that described in dioxane (Figure 11).

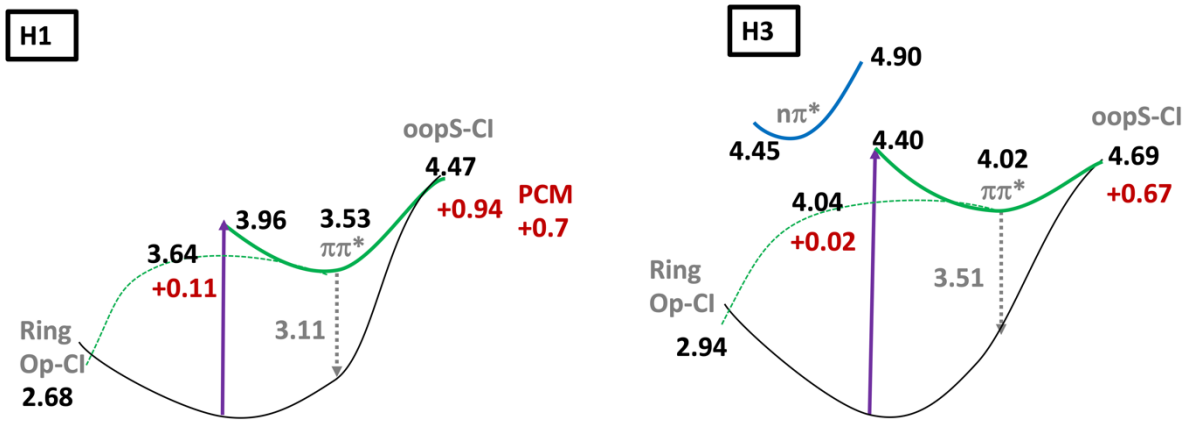

Figure 11. Potential Energy Surfaces in water. PCM/TD-M052X/6-31+G(d,p) calculations on the model containing 6 water molecules. PCM means the values obtained including only bulk solvent effects. Adiabatic Energies in black, vertical energies in gray and barriers in red (eV). The dashed green curve schematically depicts the path connecting the minima of the ππ* state with the Ring-Op-CI.

The paths leading to the RingOp and the oopS crossing regions described in the gas phase and dioxane are also present in water. The energy gaps between the ππ* minima and oopS-CI are clearly larger (+0.6 eV for H3 and 0.9 for H1) compared to gas phase (+0.3 for both tautomers), making this deactivation route very unlikely. We have also performed a test only with PCM, i.e. without water molecules, and obtained +0.7 eV. Consequently, it is likely that also in water the most accessible non-radiative decay path provides for the NS bond to elongate on the S_1 surface (bond length > 2.3 Å), before reforming after the crossing with S_0. However, the energy barrier computed in water for H1 is more than twice as large as that found in dioxane (0.11 eV and 0.17 eV with the 6-31G(d) and the aug-cc-pvtz basis sets, respectively). For the H3 tautomer, the path leading to the RingOp is easier than for H1, and the associated energy barrier is extremely small (0.02 eV with the 6-31G(d), 0.07 eV with the aug-cc-pvtz basis set) with the maximum of the barrier (Figure S12) very close to the ππ* minimum, suggesting that this path provides an effective deactivation route. Moreover, for ^{tz}G-H3, despite the destabilization of nπ* due to solute-solvent hydrogen bonds, the energy gap to ππ*1 (0.25-0.4 eV, depending on the basis set) is sufficiently small to make its population possible, if vibronic and thermal effects were considered. Finally, as discussed in the SI, the triplet electronic state is predicted to be more effectively populated for ^{tz}G-H3 than for ^{tz}G-H1 tautomer, since the Spin Orbit Coupling is, on the average, larger for the former.

Our calculations therefore suggest that the main non-radiative decay pathway of ^{tz}G-H1, namely the elongation of the NS bond in the RingOp region, is less efficient in water than in dioxane, explaining the higher ϕ value in water. This pathway is expected to be much more efficient for ^{tz}G-H3. As for the latter, the other possible non-radiative paths to non-emissive states are also more favored than for ^{tz}G-H1, since the dark nπ* state is closer to the bright one and the

coupling with triplet states larger. These factors account of the very low ϕ value measured for ^{tz}G-H3.

C. Conclusion

In this study we integrated steady-state and time-resolved fluorescence spectroscopy with QM calculations to characterize the photophysics of ^{tz}G, a promising fluorescent guanosine surrogate, in a set of four solvents of different polarity and proticity. Our results confirm that ^{tz}G has all the features of a suitable DNA/RNA fluorescent probe. Besides being perfectly isosteric with G, the absorption and emission spectra are substantially red-shifted relative to the natural nucleobases and amino acids, facilitating its selective excitation for investigating NA/protein interactions. Moreover, comparison of the data of ^{tz}G in water and methanol shows that a decrease in the H-bonding ability of the solvent results in a substantial decrease in both the fluorescence quantum yield and lifetime, whereas an even more dramatic decrease in these parameters is observed for low polarity non-hydrogen bonding solvents such as 1,4-dioxane and ethyl acetate. Therefore, ^{tz}G appears as a sensitive probe of the polarity and H-bonding ability of the embedding medium.

Furthermore, our integrated investigation revealed that ^{tz}G exhibits complex photophysics, which can be used as an additional source of information for its immediate environment. Indeed, depending on the solvent, different tautomers, identified with the help of QM calculations, can coexist at room temperature. Whilst the H1 keto-amino tautomer is dominant both in apolar aprotic and polar protic solvents, an additional enol tautomer is apparent in the former solvents and H3 keto-amino tautomer in the latter. Interestingly, all these tautomers have different spectral properties. When compared to H1, the enol tautomer shows red-shifted absorption and emission along with a long-lived fluorescence lifetime. The H3 tautomer, instead, shows blue-shifted absorption and emission, along with very low emission quantum yield ($<10^{-2}$) and short fluorescence lifetime (0.3 ns). Since the equilibria between tautomers are highly dependent on the solvent, changes in these equilibria should be observed for any investigated phenomenon that alters ^{tz}G's environment and thus, could be used as an additional information channel.

It is also interesting to compare the behavior of ^{tz}G with that of thG, which has been thoroughly investigated, both as a 'free' nucleotide and when incorporated into DNA.(10)(12)(13)(46)(21)(15)(47) While thG differs from ^{tz}G by the substitution of the N7 atom with a CH group, there are several similarities between these two probes. They notably show similar absorption and emission maxima. Moreover, for both probes, the H1 keto-amino tautomer is the most stable tautomer in polar hydrogen bonding and in non-polar aprotic solvents. Additionally, for both ^{tz}G and thG a 'blue-shifted' tautomer, H3, is populated in water, but disappears in aprotic solvents. Our study also highlights significant differences in the photophysics of ^{tz}G and thG. The most important one involves the existence of an additional non-radiative decay channel for ^{tz}G, involving the elongation of the NS bond. This deactivation route is highly impacted by solvent and responsible for the lower and more solvent-sensitive ϕ values for the H1 tautomer of ^{tz}G compared to thG. More strikingly, whilst thG-H3 exhibits ϕ similar to that of thG-H1 in water and methanol (in line with the indications of the QM calculations), the ϕ value of ^{tz}G-H3 is two orders of magnitudes smaller than that of ^{tz}G-H1 as a result of the almost barrierless pathway of the ππ* minimum to the CI or the dark nπ* state and the larger coupling with triplet states. It is thus clear that the mere substitution of the CH

group with the N atom, and therefore of the CS with the weaker NS bond, has a significant effect on the photophysics of ^{tz}G. Another difference between these two G analogues concerns the additional ^{tz}G-enol tautomer, red-shifted with respect to ^{tz}G-H1, which is instead much less stable for thG. Interestingly, ^{tz}G-enol is more emissive than ^{tz}G-H1.

Overall, although ^{tz}G is somewhat less bright than thG, its distinct deactivation pathways give it a much higher sensitivity to the polarity and H-bonding properties of its surroundings. This higher sensitivity can be seen from the values of its quantum yield and main fluorescence lifetime which drop by ~ 30% in methanol and ~75% in dioxane compared to water. In contrast, the values of the corresponding parameters of thG show only limited fluctuations between these three solvents with even the same ϕ value in water and dioxane.(21) This profound environmental sensitivity of ^{tz}G could potentially be exploited for studying guanine nucleotide-binding proteins, that act as molecular switches in numerous cellular pathways.(48) The functions of G proteins rely on their ability to bind to and hydrolyze guanosine triphosphate (GTP) to guanosine diphosphate (GDP). Therefore, being isostructural and isofunctional to G, ^{tz}G in its triphosphate form should allow a sensitive monitoring of these steps, as well as conformational changes and subsequent molecular events. Similarly, when incorporated in oligonucleotides, ^{tz}G is expected to exhibit a more sensitive response than thG upon perturbation of its immediate environment, imposed by protein binding or local conformational changes. Another factor that may contribute to a more sensitive response of ^{tz}G compared to thG when incorporated into oligonucleotides is the very low ϕ value of its H3 tautomer. Assuming that the equilibrium between the two ^{tz}G tautomers in single-stranded oligonucleotides is comparable to that of the free monomer while the H1 tautomer is strongly favoured in DNA duplexes as for thG,(14) we should observe a stronger fluorescence change upon single-stranded to double-stranded transition for ^{tz}G-labeled oligonucleotides compared to thG-labeled ones.

D. Conflicts of interest

The authors declare no conflict of interest.

E. Acknowledgements

This work and O.T. were supported by the Agence Nationale de la Recherche (ANR GQFluodynint), the Labex NIE, the Région Grand-Est (EpiRNA project) and the Centre National de la Recherche Scientifique (CNRS). Y. M. is grateful to the Institut Universitaire de France (IUF) for support and providing additional time to be dedicated to research. RI thanks financial support from CN3, National Center for Gene Therapy and Drugs based on RNA technology, funded by the European Union-NextGenerationEU-PNRR. RI also thanks the CNR program "Progetti di Ricerca @cnr", project UCATG4, and Nutrage FOE 2022 for financial support. L. M.F. thanks the CCC-UAM for generous allocation of computing time and the Madrid Government (Comunidad de Madrid-Spain) under the Multiannual Agreement with Universidad Autónoma de Madrid in the line Support to Young Researchers, in the context of the V PRICIT (Regional Programme of Research and Technological Innovation) (SI3/PJI/2021-00331).

F. References

1. Middleton, C. T., de La Harpe, K., Su, C., Law, Y. K., Crespo-Hernández, C. E. and Kohler, B. (2009) DNA Excited-State Dynamics: From Single Bases to the Double Helix.

Annu. Rev. Phys. Chem. **60**, 217–239. https://doi.org/10.1146/annurev.physchem.59.032607.093719.

2. Sinkeldam, R. W., Greco, N. J. and Tor, Y. (2010) Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **110**, 2579–2619. https://doi.org/10.1021/cr900301e.

3. Xu, W., Chan, K. M. and Kool, E. T. (2017) Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **9**, 1043–1055. https://doi.org/10.1038/nchem.2859.

4. Unruh, J. R., Gokulrangan, G., Lushington, G. H., Johnson, C. K. and Wilson, G. S. (2005) Orientational Dynamics and Dye-DNA Interactions in a Dye-Labeled DNA Aptamer. *Biophys. J.* **88**, 3455–3465. https://doi.org/10.1529/biophysj.104.054148.

5. Wang, S., Gaylord, B. S. and Bazan, G. C. (2004) Fluorescein Provides a Resonance Gate for FRET from Conjugated Polymers to DNA Intercalated Dyes. *J. Am. Chem. Soc.* **126**, 5446–5451. https://doi.org/10.1021/ja035550m.

6. Iqbal, A., Wang, L., Thompson, K. C., Lilley, D. M. and Norman, D. G. (2008) The structure of cyanine 5 terminally attached to double-stranded DNA: implications for FRET studies. *Biochemistry* **47**, 7857–7862. https://doi.org/10.1021/bi800773f.

7. Jones, A. C. and Neely, R. K. (2015) 2-aminopurine as a fluorescent probe of DNA conformation and the DNA-enzyme interface. *Q. Rev. Biophys.* **48**, 244–279. https://doi.org/10.1017/S0033583514000158.

8. Wilhelmsson, L. M. (2010) Fluorescent nucleic acid base analogues. *Quarterly reviews of biophysics* **43**, 159–183. https://doi.org/10.1017/S0033583510000090.

9. Dziuba, D., Didier, P., Ciaco, S., Barth, A., Seidel, C. A. M. and Mély, Y. (2021) Fundamental photophysics of isomorphic and expanded fluorescent nucleoside analogues. *Chem. Soc. Rev.* **50**, 7062–7107. https://doi.org/10.1039/D1CS00194A.

10. Shin, D., Sinkeldam, R. W. and Tor, Y. (2011) Emissive RNA Alphabet. *J. Am. Chem. Soc.* **133**, 14912–14915. https://doi.org/10.1021/ja206095a.

11. Rovira, A. R., Fin, A. and Tor, Y. (2015) Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **137**, 14602–14605. https://doi.org/10.1021/jacs.5b10420.

12. Park, S., Otomo, H., Zheng, L. and Sugiyama, H. (2014) Highly emissive deoxyguanosine analogue capable of direct visualization of B–Z transition. *Chemical communications* **50**, 1573–1575. https://doi.org/10.1039/C3CC48297A.

13. Sholokh, M., Sharma, R., Shin, D., Das, R., Zaporozhets, O. A., Tor, Y. and Mely, Y. (2015) Conquering 2-Aminopurine's Deficiencies: Highly Emissive Isomorphic Guanosine Surrogate Faithfully Monitors Guanosine Conformation and Dynamics in DNA. *J. Am. Chem. Soc.* **137**, 3185–3188. https://doi.org/10.1021/ja513107r.

14. Sholokh, M., Improta, R., Mori, M., Sharma, R., Kenfack, C., Shin, D., Voltz, K., Stote, R. H., Zaporozhets, O. A. , Botta, M. , Tor, Y. and Mély, Y. (2016) Tautomers of a fluorescent G surrogate and their distinct photophysics provide additional information

channels. *Angewandte Chemie International Edition* **55**, 7974–7978. https://doi.org/10.1002/anie.201601688.

15. Kuchlyan, J., Martinez-Fernandez, L., Mori, M., Gavvala, K., Ciaco, S., Boudier, C., Richert, L., Didier, P., Tor, Y., Improta, R. and Mély, Y. (2020) What Makes Thienoguanosine an Outstanding Fluorescent DNA Probe? *J Am Chem Soc* **142**, 16999–17014. https://doi.org/10.1021/jacs.0c06165.

16. Brown, J. (2020) Unraveling the structure and biological functions of RNA triple helices. *Wiley Interdisciplinary Reviews: RNA* **11**, e1598. https://doi.org/10.1002/wrna.1598.

17. Jana, J., Mohr, S., Vianney, Y. M. and Weisz, K. (2021) Structural motifs and intramolecular interactions in non-canonical G-quadruplexes. *RSC Chem. Biol.* **2**, 338–353. https://doi.org/10.1039/D0CB00211A.

18. Bucardo, M. S., Wu, Y., Ludford, P. T., Li, Y., Fin, A. and Tor, Y. (2021) Real-Time Monitoring of Human Guanine Deaminase Activity by an Emissive Guanine Analog. *ACS Chem Biol* **16**, 1208–1214. https://doi.org/10.1021/acschembio.1c00232.

19. Li, Y., Ludford III, P. T., Fin, A., Rovira, A. R. and Tor, Y. (2020) Enzymatic Syntheses and Applications of Fluorescent Cyclic Dinucleotides. *Chemistry–A European Journal* **26**, 6076–6084. https://doi.org/10.1002/chem.202001194.

20. Ciaco, S., Gavvala, K., Greiner, V., Mazzoleni, V., Didier, P., Ruff, M., Martinez-Fernandez, L., Improta, R. and Mély, Y. (2022) Thienoguanosine brightness in DNA duplexes is governed by the localization of its ππ* excitation in the lowest energy absorption band. *Methods and Applications in Fluorescence* **10**, 035003. https://doi.org/10.1088/2050-6120/ac6ab6.

21. Martinez-Fernandez, L., Gavvala, K., Sharma, R., Didier, P., Richert, L., Segarra Martì, J., Mori, M., Mely, Y. and Improta, R. (2019) Excited-State Dynamics of Thienoguanosine, an Isomorphic Highly Fluorescent Analogue of Guanosine. *Chemistry–A European Journal* **25**, 7375–7386. https://doi.org/10.1002/chem.201900677.

22. Didier, P., Kuchlyan, J., Martinez-Fernandez, L., Gosset, P., Léonard, J., Tor, Y., Improta, R. and Mély, Y. (2020) Deciphering the pH-dependence of ground-and excited-state equilibria of thienoguanine. *Physical Chemistry Chemical Physics* **22**, 7381–7391. https://doi.org/10.1039/C9CP06931C.

23. Rovira, A. R., Fin, A. and Tor, Y. (2017) Expanding a fluorescent RNA alphabet: synthesis, photophysics and utility of isothiazole-derived purine nucleoside surrogates. *Chemical science* **8**, 2983–2993. https://doi.org/10.1039/C6SC05354H.

24. Sanches de Araújo, A. V., Valverde, D., Canuto, S. and Borin, A. C. (2020) Solvation Structures and Deactivation Pathways of Luminescent Isothiazole-Derived Nucleobases: tzA, tzG, and tzI. *J. Phys. Chem. A* **124**, 6834–6844. https://doi.org/10.1021/acs.jpca.0c03398.

25. Lakowicz, J. R. (2006) Principles of fluorescence spectroscopy. Springer. https://doi.org/10.1007/978-0-387-46312-4.

26. Eaton, D. F. (1988) Reference materials for fluorescence measurement. **60**, 1107–1114. https://doi.org/10.1351/pac198860071107.

27. Muller, P.-A. Fluorescence Quantum Yields (QY) and Lifetimes (τ) for Alexa Fluor Dyes—Table 1.5. *Fluorescence Quantum Yields (QY) and Lifetimes (τ) for Alexa Fluor Dyes—Table 1.5*. Available at: https://www.thermofisher.com/fr/fr/home/references/molecular-probes-the-handbook/tables/fluorescence-quantum-yields-and-lifetimes-for-alexa-fluor-dyes.html.

28. Livesey, A. and Brochon, J. (1987) Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method. *Biophysical journal* **52**, 693–706. https://doi.org/10.1016/S0006-3495(87)83264-2.

29. Smith, D. A., McKenzie, G., Jones, A. C. and Smith, T. A. (2017) Analysis of time-correlated single photon counting data: a comparative evaluation of deterministic and probabilistic approaches. *Methods and Applications in Fluorescence* **5**, 042001. https://doi.org/10.1088/2050-6120/aa8055.

30. Knutson, J. R., Beechem, J. M. and Brand, L. (1983) Simultaneous analysis of multiple fluorescence decay curves: a global approach. *Chemical physics letters* **102**, 501–507. https://doi.org/10.1016/0009-2614(83)87454-5.

31. Beechem, J. M., Ameloot, M. and Brand, L. (1985) Global analysis of fluorescence decay surfaces: excited-state reactions. *Chemical physics letters* **120**, 466–472. https://doi.org/10.1016/0009-2614(85)85642-6.

32. Preus, S. (2013) DecayFit 1.3. *FluorTools*. Available at: http://www.fluortools.com.

33. Montgomery Jr, J. A., Frisch, M. J., Ochterski, J. W. and Petersson, G. A. (2000) A complete basis set model chemistry. VII. Use of the minimum population localization method. *The Journal of Chemical Physics* **112**, 6532–6542. https://doi.org/10.1063/1.481224.

34. Perdew, J. P., Burke, K. and Ernzerhof, M. (1997) Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* **78**, 1396–1396. https://doi.org/10.1103/PhysRevLett.78.1396.

35. Perdew, J. P., Burke, K. and Ernzerhof, M. (1996) Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

36. Zhao, Y. and Truhlar, D. G. (2008) Density functionals with broad applicability in chemistry. *Accounts of chemical research* **41**, 157–167. https://doi.org/10.1021/ar700111a.

37. Zhao, Y., Schultz, N. E. and Truhlar, D. G. (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *Journal of chemical theory and computation* **2**, 364–382. https://doi.org/10.1021/ct0502763.

38. Improta, R., Santoro, F. and Blancafort, L. (2016) Quantum Mechanical Studies on the Photophysics and the Photochemistry of Nucleic Acids and Nucleobases. *Chem. Rev.* **116**, 3540–3593. https://doi.org/10.1021/acs.chemrev.5b00444.

39. Martínez Fernández, L., Santoro, F. and Improta, R. (2022) Nucleic Acids as a Playground for the Computational Study of the Photophysics and Photochemistry of Multichromophore Assemblies. *Acc. Chem. Res.* **55**, 2077–2087. https://doi.org/10.1021/acs.accounts.2c00256.

40. Asha, H., Green, J. A., Esposito, L., Martinez-Fernandez, L., Santoro, F. and Improta, R. (2022) Effect of the Thermal Fluctuations of the Photophysics of GC and CG DNA Steps: A Computational Dynamical Study. *J. Phys. Chem. B* **126**, 10608–10621. https://doi.org/10.1021/acs.jpcb.2c05688.

41. Tomasi, J., Mennucci, B. and Cammi, R. (2005) Quantum mechanical continuum solvation models. *Chemical reviews* **105**, 2999–3094. https://doi.org/10.1021/cr9904009.

42. Miertuš, S., Scrocco, E. and Tomasi, J. (1981) Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chemical Physics* **55**, 117–129. https://doi.org/10.1016/0301-0104(81)85090-2.

43. Gao, X., Bai, S., Fazzi, D., Niehaus, T., Barbatti, M. and Thiel, W. (2017) Evaluation of spin-orbit couplings with linear-response time-dependent density functional methods. *Journal of chemical theory and computation* **13**, 515–524. https://doi.org/10.1021/acs.jctc.6b00915.

44. Avila Ferrer, F. J., Cerezo, J., Stendardo, E., Improta, R. and Santoro, F. (2013) Insights for an Accurate Comparison of Computational Data to Experimental Absorption and Emission Spectra: Beyond the Vertical Transition Approximation. *J. Chem. Theory Comput.* **9**, 2072–2082. https://doi.org/10.1021/ct301107m.

45. Karunakaran, V., Kleinermanns, K., Improta, R. and Kovalenko, S. (2009) Photoinduced dynamics of guanosine monophosphate in water from broad-band transient absorption spectroscopy and quantum-chemical calculations. *Journal of the American Chemical Society* **131**, 5839–5850. https://doi.org/10.1021/ja810092k.

46. Sholokh, M., Sharma, R., Grytsyk, N., Zaghzi, L., Postupalenko, V. Y., Dziuba, D., Barthes, N. P., Michel, B. Y., Boudier, C. , Zaporozhets, O. A., Tor, Y., Burger, A. and Mély, Y. (2018) Environmentally sensitive fluorescent nucleoside analogues for surveying dynamic interconversions of nucleic acid structures. *Chemistry–A European Journal* **24**, 13850–13861. https://doi.org/10.1002/chem.201802297.

47. Dziuba, D. (2022) Environmentally sensitive fluorescent nucleoside analogues as probes for nucleic acid–protein interactions: molecular design and biosensing applications. *Methods and Applications in Fluorescence* **10**, 044001. https://doi.org/10.1088/2050-6120/ac7bd8.

48. Simon, M. I., Strathmann, M. P. and Gautam, N. (1991) Diversity of G proteins in signal transduction. *Science* **252**, 802–808. https://doi.org/10.1126/science.1902986.

49. Andersson, K., Malmqvist, P. A., Roos, B. O., Sadlej, A. J. and Wolinski, K. (1990) Second-order perturbation theory with a CASSCF reference function. *Journal of Physical Chemistry* **94**, 5483–5488. https://doi.org/10.1021/j100377a012.

50. Andersson, K., Malmqvist, P. and Roos, B. O. (1992) Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of chemical physics* **96**, 1218–1226. https://doi.org/10.1063/1.462209.

51. Finley, J. (1998) The multi-state CASPT2 method. *Chem. Phys. Lett* **288**, 9–306. https://doi.org/10.1016/S0009-2614(98)00252-8.

52. Ghigo, G., Roos, B. O. and Malmqvist, P.-Å. (2004) A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical physics letters* **396**, 142–149. https://doi.org/10.1016/j.cplett.2004.08.032.

53. Forsberg, N. and Malmqvist, P.-Å. (1997) Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett* **274**, 196–204. https://doi.org/10.1016/S0009-2614(97)00669-6.

54. Fdez. Galván, I., Vacher, M., Alavi, A., Angeli, C., Aquilante, F., Autschbach, J., Bao, J. J., Bokarev, S. I., Bogdanov, N. A. and Carlson, R. K. (2019) OpenMolcas: From source code to insight. *Journal of chemical theory and computation* **15**, 5925–5964. http://hdl.handle.net/10044/1/74530.